融合ECA机制与DenseNet201的水稻病虫害识别方法

潘晨露，张正华*，桂文豪，马家俊，严晨曦，张晓敏

（扬州大学 信息工程学院（人工智能学院），江苏扬州 225127）

摘　要：［**目的/意义**］针对传统人工识别病虫害存在的效率过低、成本过高等问题，提出一种融合ECA（Efficient Channel Attention）注意力机制与DenseNet201的水稻图像识别模型GE-DenseNet（G-ECA DenseNet）。［**方法**］首先在ECA机制上引入Ghost模块的思想构成G-ECA Layer结构，增强其提取特征的能力。其次，在DenseNet201原有的Dense Block前引入G-ECA Layer，使模型具有更优的通道特征提取能力。由于实验所用的数据集较小，将DenseNet201在ImageNet数据集上预训练的权重参数迁移到GE-DenseNet中。训练时，采用Focal Loss函数来解决各分类样本不均衡的问题。同时，使用Adam优化器以避免在模型训练初期由于部分权重随机初始化而导致反向传播的梯度变化剧烈的问题，在一定程度上削弱了网络训练的不确定性。［**结果和讨论**］在包含水稻胡麻斑病、水稻铁甲虫、稻瘟病与健康水稻的3355张图像数据集上进行了实验测试，识别准确率达到83.52%。由GE-DenseNet模型的消融对比实验可得，引入了Focal Loss函数与G-ECA Layer层之后，模型准确率上升2.27%。将所提模型与经典NasNet（4@1056）、VGG-16和ResNet50模型相比，分类准确率分别提高了6.53%、4.83%和3.69%；相较于原始的DenseNet201，对水稻铁甲虫的识别准确率提升达20.32%。［**结论**］加入G-ECA Layer结构能够使模型更为准确地捕捉适合于水稻病虫害识别的特征信息，从而使GE-DenseNet模型能够实现对不同水稻病虫害图像更为准确地识别，为及时防治病虫害，减少各类损失提供技术支持。

关键词：DensetNet201；ECA注意力机制；病虫害识别；迁移学习；卷积神经网络；Ghost模块

中图分类号：S126　　　**文献标志码：**A　　　**文章编号：**SA202305002

引用格式：潘晨露, 张正华, 桂文豪, 马家俊, 严晨曦, 张晓敏. 融合ECA机制与DenseNet201的水稻病虫害识别方法[J]. 智慧农业(中英文), 2023, 5(2): 45-55.

PAN Chenlu, ZHANG Zhenghua, GUI Wenhao, MA Jiajun, YAN Chenxi, ZHANG Xiaomin. Rice disease and pest recognition method integrating ECA mechanism and DenseNet201[J]. Smart Agriculture, 2023, 5(2): 45-55.

1　引　言

水稻作为中国主要的粮食作物之一，2022年全国种植面积已达2945万公顷[1]。然而，病虫害对水稻产量有较大影响，轻则造成减产，重则绝收。因此，如何准确、高效地识别水稻病虫害并且及时落实对应的防治措施已成为亟待解决的问题。传统人工识别病虫害往往存在效率过低、成本过高等问题。从图像分类和图像分割[2]的角度对稻叶图像进行处理也都需要经验丰富的农业工作者或研究人员对图像进行人工标注，数据集制作过程难度大[3]。

近年来，随着人工智能的快速发展，卷积神经网络（Convolutional Neural Networks，CNN）也逐步应用于农作物病虫害的识别。Brahimi等[4]利用AlexNet将收集到的番茄叶片病害图像分类为9种病害，通过可视化的方法定位叶片中的病害区域，

收稿日期：2023-05-07

基金项目：2020年江苏省现代农业发展项目（2020-SJ-003-YD03）；扬州大学学科特区学科交叉课题（yzuxk202008）；2022年江苏省大学生创新训练计划项目重点项目（国家级）（202211117065Z）

作者简介：潘晨露，研究方向为计算机视觉。E-mail：2300935324@qq.com

*通信作者：张正华，教授，研究方向为智慧农业、实时信号处理。E-mail：zhangzh@yzu.edu.cn

证明AlexNet具有较好的识别效果。Rahman等[5]对VGG-16和Inception-V3进行了模型微调，并提出了一种两阶段的小型CNN架构，在自建的1426张水稻病害图像的数据集中得到了93.3%的平均准确率。李淼等[6]在AlexNet和VGG（Visual Geometry Group）网络模型的基础上，引入批归一化算法，改善了模型的过拟合问题，在由3种黄瓜病害图像和5种水稻病害图像构成的自制数据集AES-IMAGE上，平均识别准确率分别达到了95.93%和95.42%。Sharma等[7]引入一系列集合块用以提取深层特征，同时采用可分离卷积来减小计算量，提出了DLMC-Net模型，在四个公开的柑橘、黄瓜、葡萄和番茄的叶片数据集中，平均准确率分别达到93.56%、92.34%、99.50%和96.56%。Monu等[8]在VGG-16模型基础上，引入局部二值特征进行融合，并采用随机森林法分类，有效提升了模型的鲁棒性，在公开的PlantVillage甜椒叶片数据集上，准确率达到99.75%。王东方和汪军[9]对SE-ResNeXt-101模型进行改进，结合迁移学习，提出了一种不指定农作物种类的病害分类模型TL-SE-ResNeXt-101，在重构的AI Challenger 2018农作物病害数据集上测试平均准确率达到98%。

以上网络模型通常只从调整深度或宽度两方面入手，不能有效利用通道特征信息进行特征提取。本研究针对水稻胡麻斑病（BrownSpot）、水稻铁甲虫（Hispa）、稻瘟病（LeafBlast）这3种常见病虫害识别问题，通过增加模型的通道特征表达能力，提出了一种融合高效通道注意力（Efficient Channel Attention，ECA）机制与DenseNet201的水稻病虫害识别方法。利用该方法可以有效对水稻叶片图像进行分析，结合实时监测数据构建水稻病虫害检测与预警系统，有助于在病虫害发生早期阶段及时采取防治措施，减少各类损失。同时，收集大量水稻病虫害分析数据也可以促进病虫害研究与防治技术的不断改进和优化。

2 实验材料

2.1 水稻病虫害图像采集

选取水稻胡麻斑病、水稻铁甲虫和稻瘟病这3种水稻常见的病虫害，将其与健康水稻的叶片图像构成数据集，并进行病虫害图像识别。实验数据采集点为江苏省仪征市壮禾米业专业合作社，为使数据符合实际应用场景，采用SONY DSC-W800相机每隔2 h手动从水稻顶部、侧面等多角度进行拍摄，从而获取不同光照条件、不同视角、不同遮挡环境下水稻叶片的图像1250张。另外，在Kaggle数据库（https://www.kaggle.com/）收集有关病虫害样本。数据集总计3355张图像，其中健康叶片1488张，水稻胡麻斑病523张，水稻铁甲虫565张，稻瘟病779张。

2.2 水稻病虫害图像预处理

由于病虫害数据的原始特征较为接近，本研究首先按照9∶1的比例将数据集划分成为训练集与测试集，再对训练集进行数据增强[10]，从而获得更大的训练集分配，提高网络模型的泛化能力，防止出现过拟合的情况。使用随机选择一个感兴趣区域（Region of Interest，ROI）的方式实现数据集样本图像局部比例为1.1～1.25缩放，从而模拟实际拍摄过程中由于植物距离摄像头远近不同，只有部分叶片被拍摄到的情况。具体步骤包括：在图像左上角20×20像素内随机选取ROI矩形的起始坐标点，再选取180～200之间的任意数作为ROI的边长，从而形成一个ROI区域。另外，使用随机旋转一定角度后进行裁切的方式模拟叶片不同角度拍摄图像，具体步骤包括：（1）随机选取一个旋转角度；（2）针对图像的中心点和得到的旋转角度生成相应的仿射变换矩阵，实现图像的旋转操作；（3）通过选取最大内接矩形的方法去除黑边。

最终增强后训练集大小为18,018张，不同尺寸的图像统一转换像素为224×224×3，其中3代表彩色图像中RGB三个通道的信息，以便模型训练，样本集具体构成如表1所示。

数据样本展示如图1所示。水稻胡麻斑病（图1（b））的显著特征是其叶片上分布有大小不一的椭圆形褐色病斑，病斑边缘明显并有黄色晕圈环绕。而水稻铁甲虫（图1（c））则表现为叶片上有形状、大小各异的被铁甲虫啃食后的缺口以及幼虫取食果肉后叶片表面呈现的枯白色斑痕。稻瘟病

表 1 水稻病虫害数据集分布表

Table1 Distribution of rice pest and disease dataset

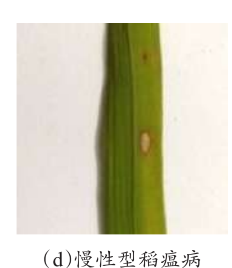
数据类别	增强后训练集/幅	测试集/幅
健康叶片	7968	160
水稻胡麻斑病	2754	64
水稻铁甲虫	3006	64
稻瘟病	4290	64
合计	18,018	352

数据集中样本特征分为慢性型和白斑型这两种症状，慢性型（图1（d））具体表现为病斑呈菱形或纺锤形，最外层为黄色晕圈，内环为褐色，中央为灰白色；白斑型（图1（e））表现为圆形或不规则形状的白色或灰白色斑点。


(a)健康水稻


(b)水稻胡麻斑病


(c)水稻铁甲虫


(d)慢性型稻瘟病


(e)白斑型稻瘟病

图 1 水稻病虫害样本图像

Fig. 1 Images of rice pest and disease samples

3 水稻病虫害识别模型构建

为了获得更丰富的特征信息，本研究选用DenseNet201 [11] 模型作为基准网络。为使模型在获取不同层间关联性的同时更高效地获取通道信息，引入了基于Ghost思想改进的ECA模块，将其命名为G-ECA Layer（Ghost-ECA Layer），并利用这一结构优化DenseNet201网络，提出了一种融合注意力机制与DenseNet201的水稻病虫害识别方法GE-DenseNet（G-ECA DenseNet）。

3.1 G-ECA Layer的构建

本研究基于Ghost模块[12] 的思想，改变了ECA注意力机制[13] 的设计，提出G-ECA Layer这一结构。具体设计如图2所示。

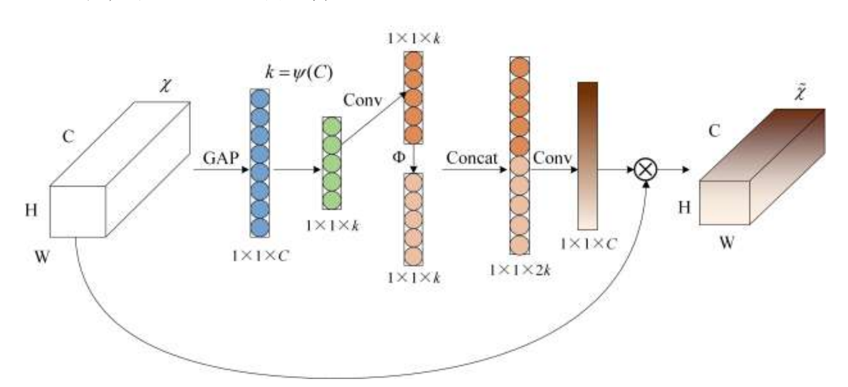


图 2 G-ECA Layer结构图

Fig.2 Structure of G-ECA Layer

假设输入特征图的尺寸为$H\times W\times C$，其中H代表特征图的高度，W代表特征图的宽度，C代表特征图的通道数。首先使用全局平均池化（Global Average Pooling，GAP）[14] 对其进行处理，将其压缩为$1\times1\times C$的特征图。随后，将压缩后的特征图，通过卷积得到大小为$1\times1\times k$的特征图。

卷积核[15] 数量可以通过对通道维数进行非线性映射从而实现自适应确定，卷积核适应函数定义如公式（1）。

$$k=\psi(C)=\left|\frac{\log_2(C)}{\gamma}+\frac{b}{\gamma}\right|_{odd} \tag{1}$$

其中，k表示卷积核的数量（只能取奇数）；C

表示通道数；γ和b用于改变通道数C与卷积核数量之间的比例。通过该函数，G-ECA Layer结构可以实现：在通道数较大的层，使用较多卷积核，加强对输入通道特征信息的提取；在通道数较小的层，使用较少的卷积核，减少跨通道交互从而减少计算量。

卷积后得到的大小为1×1×k的特征图，第一部分使用大小为1的卷积核进行常规卷积操作，得到本征特征图$Y \in R^{1 \times 1 \times k}$，卷积运算可以表示为公式（2）。

$$Y = X * f' + g \tag{2}$$

其中，$*$为卷积运算的符号；g为偏置项；$f' \in R^{k \times 1 \times 1 \times k}$代表大小为1的卷积核，其中输出通道数等于输入通道数，对原始特征进行必要的特征提取。

第二部分则利用已经生成的本征特征图Y进行简单的线性变换Φ，得到Ghost特征图，如公式（3）。

$$y'_i = \Phi(y_i),\ (i = 1, ..., n) \tag{3}$$

其中，y_i是Y中第i个本征特征图；Φ是对本征特征图进行的线性运算；y_i'表示第 i 个原始特征图经过变换后的Ghost特征图。

最后，将两部分特征图通过拼接（Concatenation）操作[16]叠加，得到输出为1×1×2k的特征图。通过1×1的卷积进行通道特征学习，得到1×1×C的通道注意力特征图。再将该特征图与原始输入的H×W×C的特征图按照通道顺序进行逐通道点乘操作，从而输出与原始输入特征图尺寸一致的通道注意力特征图。

G-ECA Layer结构，采用了Ghost思想，以线性变换代替卷积运算，对通道特征进行高效融合的同时避免了学习通道注意力信息时的维度减缩，有效提升了识别准确率。

3.2 GE-DenseNet模型构建

DenseNet201的模型结构如图3所示，模型整体由四个Dense Block组成，并使用Transition Layer层作为连接。

其中，Transition Layer层包含了一个1×1的卷积和2×2的平均池化操作，整合上一个Dense Block获取的特征同时缩小上一个Dense Block输出的尺寸。

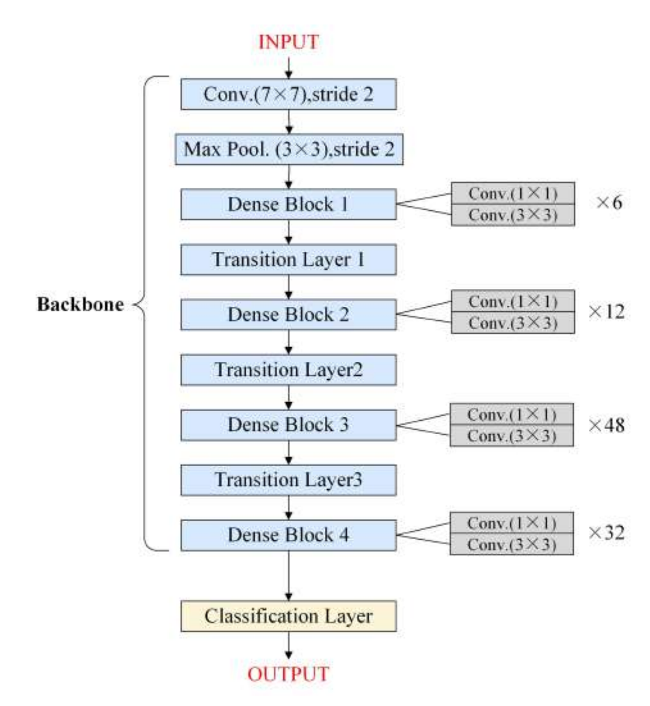


图3　DenseNet201的分层体系结构图

Fig.3　Layered architecture of DenseNet201

1×1和3×3卷积对构成了Dense Block的基本组成部分。其不仅降低了特征的维度，同时也增加了网络的非线性度，防止模型过拟合。在DenseNet201的四个Dense Block模块中，1×1和3×3卷积对的叠加数量依次为6个、12个、48个和32个，其叠加的具体构成方式如下：

假设通过前馈网络模型的特征层x_0，经过一个L层的网络模型。在传统的卷积神经网络中，其中第i层的非线性变换记为$H_i(*)$，第i层的特征层输出记为x_i，则有公式（4）：

$$x_i = H_i(x_0) \tag{4}$$

为进一步优化信息的利用效率，DenseNet引入Dense Block模块，使得L层的网络模型中第i层的输入与之前所有层的输出相关，即公式（5）：

$$x_i = H_i([x_0, x_1, \cdots, x_{i-1}]) \tag{5}$$

其中，$[x_0,\ x_1,\ \cdots,\ x_{i-1}]$是特征层的拼接操作，即将当前层之前的所有特征层输出按照通道维度叠加在一起，而非线性变换H_i，即上文中提及的1×1和3×3卷积对。由于Dense Block中每一个卷积对之间都存在特征向量进行连接，从而改善了整

个网络的信息密集度和梯度传递的信息流，可以使模型具有更好的特征利用效率。在进行梯度反向传播的计算时，每一层网络都可以接收到该层后面特征层传递的梯度信息，一定程度上减少了由于网络深度增加导致的梯度消失问题。此外，更为密集的连接在深度较大的模型结构中通常具有正则化效应，可以有效防止模型在训练过程中出现过拟合现象。

基于3.1节所述G-ECA Layer结构，本研究对DenseNet201进行了修改。如图3所示，DenseNet201整体由四个Dense Block组成，使用Transition Layer结构作为连接。但是，由于原始的Dense Block只考虑到不同层之间的关联性，忽略了图像识别过程中重要通道信息的提取，一定程度上影响了模型对水稻病虫害的识别准确率。为了优化模型识别性能，本研究以DensNet201模型为基础网络，在原有的Dense Block前引入G-ECA Layer结构，对输入给Dense Block的信息进行通道注意力的增强，以达到进一步提升识别模型准确率的目的。所提出的GE-DenseNet网络结构示意图如图4所示。

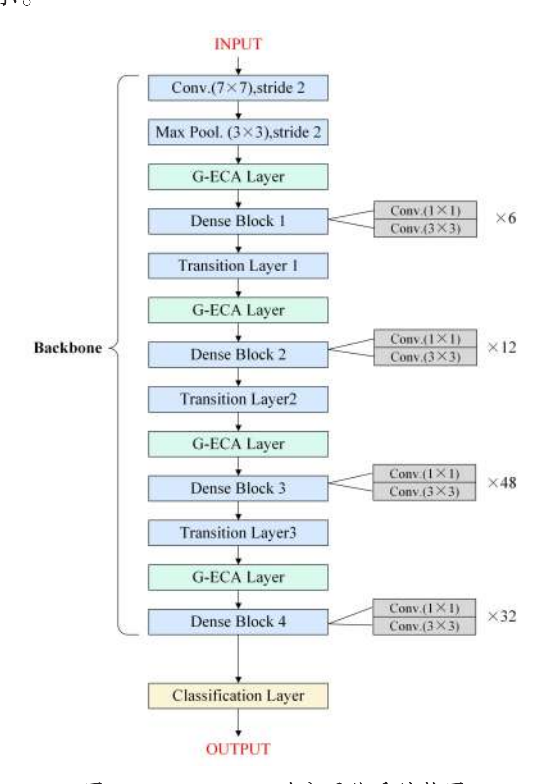


图4 GE-DenseNet的分层体系结构图

Fig. 4 Layered architecture of GE-DenseNet

3.3 Focal Loss损失函数

传统网络通常使用交叉熵损失函数（Cross Entropy Loss，CE）进行模型训练，如公式（6）。

$$CE(p,y)=\begin{cases}-\log(p) & if y=1\\ -\log(1-p) & otherwise\end{cases} \tag{6}$$

其中，CE（）表示交叉熵损失函数；p代表样本的预测值，范围为［0，1］；y代表样本的真实值，为0或1。

交叉熵损失函数是对各个训练样本交叉熵的直接求和，即各个样本的权重是一样的。然而，由于样本的类别不均衡，交叉熵损失函数识别准确率难以提升[17]。针对该问题，本研究引入了Focal Loss[18]代替交叉熵损失函数，具体为公式（7）。

$$FL(p_t)=-(1-p_t)^{\gamma}\log(p_t) \tag{7}$$

其中，p_t为某个类别的类别分数；γ为聚焦参数；$(1-p_t)^{\gamma}$为调制系数。

根据公式（7）可知：首先，当样本被错误分类时，类别分数p_t很小，$(1-p_t)$接近1，从而损失不会受到影响。而当p_t趋近于1时，$(1-p_t)$趋向于0，这意味着分类效果良好的样本的权重会被降低，因此它们对总体损失的贡献很小。

其次，当γ等于0时，Focal Loss即等同于交叉熵损失函数。随着聚焦参数γ的增加，调制系数$(1-p_t)^{\gamma}$也会增加。因此，通过调节γ能够平滑地调整易分类样本被降低权重的比例。基于这些参数，Focal Loss即可通过增加误分类样本的重要性均衡样本，从而提升网络的训练速度和精度。

3.4 Adam优化器

自适应矩估计（Adaptive Moment Estimation，Adam）[19]是一种自适应动量的随机优化方法，它在随机梯度下降方法（Stochastic Gradient Descent，SGD）的基础上引入了动量和自适应学习率的思想，用以加速梯度下降的过程，具体如公式（8）～（12）所示：

$$m_t=\beta_1 m_{t-1}+(1-\beta_1)g_t \tag{8}$$

$$v_t=\beta_2 v_{t-1}+(1-\beta_2)g_t^2 \tag{9}$$

$$\hat{m}_t=\frac{m_t}{1-\beta_1^t} \tag{10}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^{\ t}} \tag{11}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon}\hat{m}_t \tag{12}$$

公式（8）是一阶矩（动量）的计算公式，其中，m_t表示第t个训练批次的动量；g_t表示第t个训练批次的梯度。公式（8）通过超参数β_1控制历史梯度的加权平均，从而可以得到一个较为平滑的梯度值。公式（9）为二阶矩的计算公式，其中，v_t表示第t个训练批次的二阶矩估计，β_2超参数则控制历史梯度的平方和加权平均。公式（10）和（11）中，$\hat{m}_t$和$\hat{v}_t$分别为β_1与β_2对m_t与v_t的偏差矫正，用于解决训练初期由于历史数据较少引起的偏差。公式（12）为网络模型参数的更新公式，其中θ_t是第t个训练批次的参数，η是模型训练的学习率（Learning Rate），ε是一个很小的常数。将公式中的$\frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon}$看作一个整体，即表示通过历史梯度的平方和加权平均，实现学习率的动态调整。

本研究使用Adam自适应学习率梯度下降算法进行训练，可以避免在模型训练初期由于部分权重随机初始化而导致反向传播的梯度变化剧烈的问题，在一定程度上削弱了网络训练的不确定性。

4　结果比较与分析

4.1　实验环境搭建与模型训练

实验平台搭载Intel i9-12900K处理器，内存64 GB，显卡为NVIDIA RTX A4000，软件环境为Windows10 LTSC 64位操作系统，使用Anaconda虚拟环境，其中Python版本为3.8.15，深度学习框架Tensorflow版本为2.6.0-GPU、深度学习加速库为CUDA Toolkit11.3.1和cuDNN8.2.1。

GE-DenseNet模型以数据增强后的18,018张水稻病虫害图像进行训练，352张图像进行测试。同时，以VGG-16[20]、ResNet50[21]、NasNet（4@1056）[22]、DenseNet201作为对照组实验。训练时，对照组学习率lr设置为10^{-4}，采用交叉熵损失函数和Adam优化器。在训练过程中，BatchSize大小设置为32，迭代次数（Epoch）设置为50，保存训练集准确率最高的模型权重用于测试集评估。

4.2　迁移学习

迁移学习（Transform Learning）[23]是指将已经学习过相关任务的知识（参数文件、模型权重等）应用于新的相关任务中以得到更好训练效果的一种方法。

本研究的数据集分布较为集中，类别较少，迁移学习高质量数据集预训练产生的模型进行微调训练可以有效解决过拟合等问题。因此，在构建完成GE-DenseNet后，选择利用在ImageNet数据集上预训练的DenseNet201模型权重对GE-DenseNet模型中与DenseNet201模型结构类似的部分网络进行初始化，使模型获得较为丰富的先验知识，再对GE-DenseNet模型在自制数据集上进行训练。评价指标Macro Avg P表示精确率，即样本中正确样本被预测正确占所有被预测为正确样本的比率；Macro Avg R表示召回率，即样本中正确样本被预测正确占所有正确样本的比率；Macro Avg F_1为精确率和召回率的调和平均数[24]；Acc表示准确率，即判定为正确的样本实际正确和被判定为错误的样本实际错误的两类占比。

如表2所示，经过迁移学习后的GE-DenseNet模型在Acc、Macro Avg P、Macro Avg R和Macro Avg F_1各项评价指标上均提升了至少3%。由此可得，迁移学习应用于本研究所提模型下能够有效提高识别准确度，增强模型性能。

表2　GE-DenseNet模型迁移学习前后性能对比

Table 2　Performance comparison of GE-DenseNet model before and after migration learning

Model	Acc	Macro Avg P	Macro Avg R	Macro Avg F_1
GE-DenseNet	0.7983	0.8138	0.7508	0.7755
GE-DenseNet+ Transform Learning	0.8352	0.8451	0.7992	0.8169

4.3　GE-DenseNet的识别性能测试

为验证GE-DenseNet相较于DenseNet201的优越性，在实验数据集上将其二者的训练损失率和准确率进行比较，具体对比结果如图5和图6所示。

从图5可知，随着训练次数的增加，GE-

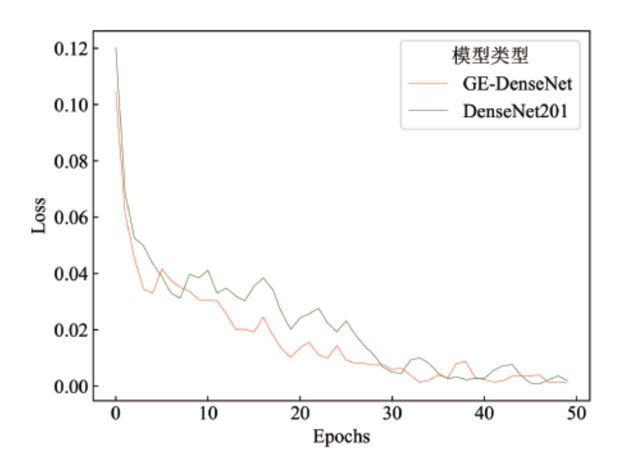


图 5　GE-DenseNet模型与DenseNet201模型的损失率变化曲线图

Fig.5　Loss rate variation curves of GE-DenseNet model and DenseNet201 model

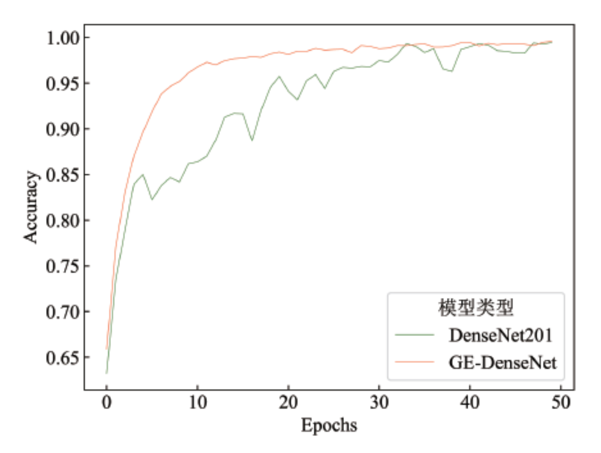
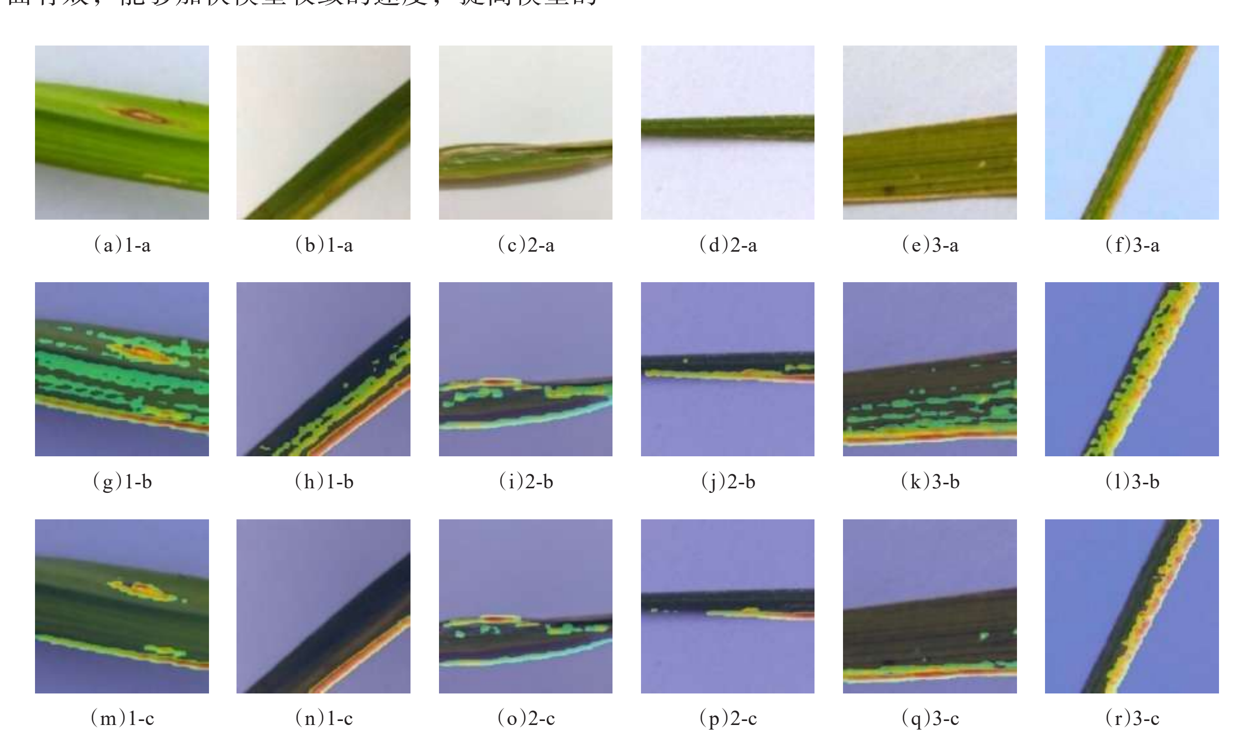


图 6　GE-DenseNet模型与DenseNet201模型的准确率变化曲线图

Fig.6　Accuracy variation curves of GE-DenseNet model and DenseNet201 model

DenseNet和DenseNet201的损失率曲线均不断下降，但GE-DenseNet曲线下降速度更快，这说明其具有更好的学习能力。从图6可知，当Epoch为15时，使用G-ECA Layer和Focal Loss的网络模型准确率快速稳定，而原始模型仍处于准确率不断上升的过程。实验证明，本研究所提方法在训练稳定性方面有效，能够加快模型收敛的速度，提高模型的稳定性，使之具有更好的识别效果。

对比图7中每一组的DenseNet201处理结果和GE-DenseNet处理结果可以发现，加入G-ECA Layer结构后，特征更密集地体现在病虫害的周围，进一步说明了G-ECA Layer结构在本研究所提模型上的有效性。



注：1为稻瘟病；2为水稻铁甲虫；3为水稻胡麻斑病；a表示输入图像；b为DenseNet201处理结果；c为GE-DenseNet处理结果

图 7　GE-DenseNet与DenseNet201模型特征层可视化结果对比图

Fig.7　Feature layer visualization results comparison of GE-DenseNet and DenseNet201 model

4.4 GE-DenseNet的消融试验

为证明本研究所提方法中各个模块的有效性，设计以下4种消融实验方法进行比较分析：

（1）使用原始的DenseNet201进行模型训练；

（2）在（1）的基础上采用Focal Loss损失函数；

（3）在（1）的基础上在每个Dense Block前引入G-ECA Layer结构；

（4）在（3）的基础上引入Focal Loss损失函数，形成本研究所提方法。

上述4种消融实验方法在本研究使用的水稻病虫害数据集上的识别结果如表3所示。

表3 GE-DenseNet模型消融试验结果

Table3 GE-DenseNet Model ablation test results

Model	Acc	Macro Avg P	Macro Avg R	Macro Avg F_1
DenseNet201	0.8125	0.8354	0.7633	0.7850
DenseNet201+ Focal	0.8239	0.8379	0.7789	0.8026
DenseNet201+ G-ECA	0.8324	0.8426	0.8047	0.8183
DenseNet201+ G-ECA +Focal	0.8352	0.8451	0.7992	0.8169

从方法（1）和方法（2）可知，引入Focal Loss损失函数代替原模型使用的交叉熵损失函数，使模型的准确率上升1.14%，从而验证了Focal Loss在数据样本不均衡的情况下对模型性能提升的有效性。从方法（1）和方法（3）可知，引入G-ECA Layer层，使模型的准确率上升1.99%，并使得模型的召回率提升4.14%，证明了加入改进后的注意力机制能够有效利用通道信息，实现更加准确地识别。进一步比较方法（1）与方法（4），可知引入了Focal Loss多分类损失与G-ECA Layer层之后，模型准确率上升2.27%，模型的召回率和F_1 Score均有较大幅度提升，体现了所提模型的有效性。

4.5 与经典深度卷积神经网络模型性能对比

将本研究所提方法在相同训练环境与相同梯度下降函数的条件下，与VGG-16、ResNet50、Nas-Net（4@1056）、DenseNet201等经典深度卷积神经网络进行性能对比。

由表4可知，GE-DenseNet相较于改进前的DenseNet201模型在稻瘟病和健康水稻的识别效果上略有下降，但其对水稻铁甲虫的识别准确率提升达20.32%，整体识别准确率也优于其他模型。这一结果表明在DenseNet201模型中加入的G-ECA Layer结构能够更为准确地捕捉适合于该应用场景的特征信息从而提高识别准确率。

表4 水稻病虫害识别实验不同模型各分类预测准确率对比

Table 4 Prediction accuracy comparison of different models for each classification in rice pest identification experiments

Model	BrownSpot Acc	Hispa Acc	LeafBlast Acc	Health Acc	Acc
VGG-16	0.8438	0.6875	0.7812	0.8063	0.7869
ResNet50	0.6875	0.5625	0.8125	0.9313	0.7983
NasNet(4@1056)	0.7500	0.5156	0.7969	0.8688	0.7699
DenseNet201	0.7500	0.5156	0.8438	0.9437	0.8125
GE-DenseNet	**0.7500**	**0.7188**	**0.7969**	**0.9313**	**0.8352**

由表5可知，本研究所提模型在各项指标上都优于其他4类经典模型，在Macro Avg R指标的比较中，DenseNet201小于VGG-16模型，但GE-DenseNet性能优于VGG-16，这表明改进后的模型在一定程度上提升了水稻病虫害识别的效果。

表5 水稻病虫害识别实验不同模型性能对比

Table 5 Performance comparison of different models for rice pest identification experiments

Model	Acc	Macro Avg P	Macro Avg R	Macro Avg F_1	Size/MB	Calculation Volume/GFLOPs
VGG-16	0.7869	0.7761	0.7797	0.7747	154.17	30.76
ResNet50	0.7983	0.8163	0.7484	0.7725	482.02	7.93
NasNet(4@1056)	0.7699	0.7647	0.7328	0.7429	218.45	1.24
DenseNet201	0.8125	0.8354	0.7633	0.7850	437.43	8.77
GE-DenseNet	**0.8352**	**0.8451**	**0.7992**	**0.8169**	**437.68**	**8.78**

通过图8可知，在训练初期，DenseNet201得益于其较强的学习能力，识别准确率增长最快，最先达到70%。而在迭代25次左右后，由于VGG-16模型结构较为简单，模型最先收敛，识别准确率为78%左右。GE-DenseNet在前40轮训练中学习速率相较VGG-16略有逊色，但由于加入了注意力机制，能够更全面地学习数据集的特征，所以在其他模型趋于稳定的情况下仍处于震荡向上的阶段，经过50次迭代后，达到了最优准确率，体现了本研究所提方法的有效性。

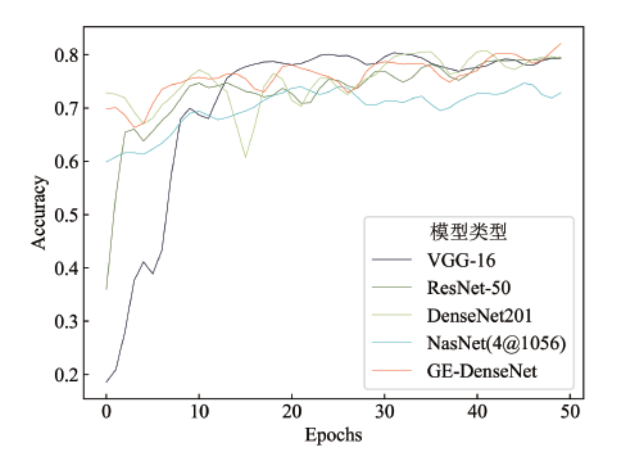


图8 不同模型识别水稻病虫害准确率变化曲线图

Fig.8 Accuracy variation curves of different models to identify rice pests and diseases

5 结 论

针对水稻病虫害图像的特点，本研究提出了一种融合ECA注意力机制与DenseNet201的图像分类模型GE-DenseNet，用以识别水稻胡麻斑病、水稻铁甲虫、稻瘟病和健康水稻4类水稻叶片图像。GE-DenseNet在DenseNet201每个Dense Block前添加利用Ghost思想改进的ECA注意力机制，从而将更多通道信息融入特征图中以提高模型的准确率。训练时，使用Focal Loss代替交叉熵函数以改善训练集样本不均衡的问题。

本研究所提方法在测试集上准确率为83.52%，比DenseNet201提高了2.27%，比经典的深度卷积神经网络NasNet（4@1056）、VGG-16和ResNet50提高了6.53%、4.83%和3.69%。综上，GE-DenseNet能够提高网络的特征提取能力，从而提高了分类准确率。

但将该模型投入实际应用仍存在一些不足之处，显著表现为模型所占内存较大不易于移动终端和硬件平台的部署。后续将尝试通过知识蒸馏、剪枝等操作对模型进行轻量化，在不显著降低识别准确率的前提下，减小模型体积，从而部署于无人机、小车和各类分布式图像检测边缘设备，方便农户进行农田的实时检测，有效减少人力成本并提高检测效率。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 徐春春, 纪龙, 陈中督, 等. 2022年我国水稻产业发展分析及2023年展望[J]. 中国稻米, 2023, 29(2): 1-4.
XU C C, JI L, CHEN Z D, et al. Analysis of China's rice industry in 2022 and the outlook for 2023[J]. China rice, 2023, 29(2): 1-4.

[2] 刘震, 纪明姝, 郭志顶, 等. 图像识别技术在病虫害防治方面的应用与展望[J]. 沧州师范学院学报, 2022, 38(1): 119-123.
LIU Z, JI M M, GUO Z D, et al. Application and prospect of image recognition technology in pest control[J]. Journal of Cangzhou normal university, 2022, 38(1): 119-123.

[3] 朱成宇. 基于轻量级卷积神经网络的水稻病害识别[D]. 南宁: 广西大学, 2022.
ZHU C Y. Rice desease recognition based on lightweight convolutional neural network[D]. Nanning: Guangxi University, 2022.

[4] BRAHIMI M, BOUKHALFA K, MOUSSAOUI A. Deep learning for tomato diseases: Classification and symptoms visualization[J]. Applied artificial intelligence, 2017, 31(4): 299-315.

[5] RAHMAN C R, ARKO P S, ALI M E, et al. Identification and recognition of rice diseases and pests using convolutional neural networks[J]. Biosystems engineering, 2020, 194: 112-120.

[6] 李淼, 王敬贤, 李华龙, 等. 基于CNN和迁移学习的农作物病害识别方法研究[J]. 智慧农业, 2019, 1(3): 46-55.
LI M, WANG J X, LI H L, et al. Method for identifying crop disease based on CNN and transfer learning[J]. Smart agriculture, 2019, 1(3): 46-55.

[7] SHARMA V, TRIPATHI A K, MITTAL H. DLMC-Net: Deeper lightweight multi-class classification model for plant leaf disease detection[J]. Ecological informatics, 2023, 75: ID 102025.

[8] MONU B, DILIP K, SUNIL K. Bell pepper leaf disease classification with LBP and VGG-16 based fused features and RF classifier[J]. International journal of information technology, 2022, 15(1): 465-475.

[9] 王东方, 汪军. 基于迁移学习和残差网络的农作物病害分类[J]. 农业工程学报, 2021, 37(4): 199-207.
WANG D F, WANG J. Crop disease classification with transfer learning and residual networks[J]. Transactions of the Chinese society of agricultural engineering, 2021, 37 (4): 199-207.

[10] 高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(8): 62-65.
GAO Y W, ZHOU B J, HU X F. Research on image recognition of convolution neural network based on data enhancement[J]. Computer technology and development, 2018, 28(8): 62-65.

[11] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 2261-2269.

[12] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 1577-1586.

[13] WANG Q L, WU B G, ZHU P F, et al. ECA-net: Efficient channel attention for deep convolutional neural networks[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 11531-11539.

[14] 张善文, 张晴晴, 李萍. 基于改进深度卷积神经网络的苹果病害识别[J]. 林业工程学报, 2019, 4(4): 107-112.
ZHANG S W, ZHANG Q Q, LI P. Apple disease identification based on improved deep convolutional neural network[J]. Journal of forestry engineering, 2019, 4(4): 107-112.

[15] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. Chinese journal of computers, 2017, 40 (6): 1229-1251.

[16] 罗大为, 方建军, 刘艳霞. 基于通道域注意力机制的特征融合方式[J]. 东北师大学报(自然科学版), 2021, 53(3): 44-48.
LUO D W, FANG J J, LIU Y X. Feature fusion methods based on channel domain attention mechanism[J]. Journal of northeast normal university (natural science edition), 2021, 53(3): 44-48.

[17] 杨莲, 石宝峰. 基于Focal Loss修正交叉熵损失函数的信用风险评价模型及实证[J]. 中国管理科学, 2022, 30 (5): 65-75.
YANG L, SHI B F. Credit risk evaluation model and empirical research based on focal loss modified cross-entropy loss function[J]. Chinese journal of management science, 2022, 30(5): 65-75.

[18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 2999-3007.

[19] 卫雅娜, 王志彬, 乔晓军, 等. 基于注意力机制与EfficientNet的轻量化水稻病害识别方法[J]. 中国农机化学报, 2022, 43(11): 172-181.
WEI Y N, WANG Z B, QIAO X J, et al. Lightweight rice disease identification method based on attention mechanism and EfficientNet[J]. Journal of Chinese agricultural mechanization, 2022, 43(11): 172-181.

[20] SENGUPTA A, YE Y T, WANG R, et al. Going deeper in spiking neural networks: VGG and residual architectures[EB/OL]. arXiv: 1802.02627, 2018.

[21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.

[22] ZOPH B, VASUDEVAN V, SHLENS J, et al. Learning transferable architectures for scalable image recognition[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 8697-8710.

[23] 黎英, 宋佩华. 迁移学习在医学图像分类中的研究进展[J]. 中国图象图形学报, 2022, 27(3): 672-686.
LI Y, SONG P H. Review of transfer learning in medical image classification[J]. Journal of image and graphics, 2022, 27(3): 672-686.

[24] 黄浩, 徐海华, 王羡慧, 等. 自动发音错误检测中基于最大化F_1值准则的区分性特征补偿训练算法[J]. 电子学报, 2015, 43(7): 1294-1299.
HUANG H, XU H H, WANG X H, et al. Maximum F_1 score criterion based discriminative feature compensation training algorithm for automatic mispronunciation detection[J]. Acta electronica sinica, 2015, 43(7): 1294-1299.

Rice Disease and Pest Recognition Method Integrating ECA Mechanism and DenseNet201

PAN Chenlu, ZHANG Zhenghua*, GUI Wenhao, MA Jiajun, YAN Chenxi, ZHANG Xiaomin

(*College of information engineering* (*College of Artificial Intelligence*), *Yangzhou University*, *Yangzhou* 225127, *China*)

Abstract:

[Objective] To address the problems of low efficiency and high cost of traditional manual identification of pests and diseases, improve the automatic recognition of pests and diseases by introducing advanced technical means, and provide feasible technical solutions for agricultural pest and disease monitoring and prevention and control, a rice image recognition model GE-DenseNet (G-ECA DenseNet) based on improved ECA (Efficient Channel Attention) mechanism with DenseNet201 was proposed.

[Methods] The leaf images of three pests and diseases, namely, brownspot, hispa, leafblast and healthy rice were selected as experimental materials. The images were captured at the Zhuanghe Rice Professional Cooperative in Yizheng, Jiangsu Province, and the camera was used to manually take pictures from multiple angles such as the top and side of rice every 2 h, thus acquiring 1250 images of rice leaves under different lighting conditions, different perspectives, and different shading environments. In addition, samples about pests and diseases were collected in the Kaggle database. There were 1488 healthy leaves, 523 images of brownspot, 565 images of hispa, and 779 images of leafblast in the dataset. Since the original features of the pest and disease data were relatively close, firstly, the dataset was divided into a training set and a test set according to the ratio of 9:1, and then data enhancement was performed on the training set. A region of interest (ROI) was randomly selected to achieve a local scale of 1.1 to 1.25 for the sample images of the dataset, thus simulating the situation that only part of the leaves were captured in the actual shooting process due to the different distance of the plants from the camera. In addition, a random rotation of a certain angle was used to crop the image to simulate the different angles of the leaves. Finally, the experimental training set contains 18,018 images and the test set contains 352 images. The GE-DenseNet model firstly introduces the idea of Ghost module on the ECA attention mechanism to constitute the G-ECA Layer structure, which replaces the convolution operation with linear transformation to perform efficient fusion of channel features while avoiding dimensionality reduction when learning channel attention information and effectively enhancing its ability to extract features. Secondly, since the original Dense Block only considered the correlation between different layers and ignores the extraction of important channel information in the image recognition process, introducing G-ECA Layer before the original Dense Block of DenseNet201 gives the model a better channel feature extraction capability and thus improved the recognition accuracy. Due to the small dataset used in the experiment, the weight parameters of DenseNet201 pre-trained on the ImageNet dataset were migrated to GE-DenseNet. During the training process, the BatchSize size was set to 32, the number of iterations (Epoch) was set to 50, and the Focal Loss function was used to solve the problem of unbalanced samples for each classification. Meanwhile, the adaptive moment estimation (Adam) optimizer was used to avoid the problem of drastic gradient changes in back propagation due to random initialization of some weights at the early stage of model training, which weakened the uncertainty of network training to a certain extent.

[Results and Discussions] Experimental tests were conducted on a homemade dataset of rice pests and diseases, and the recognition accuracy reached 83.52%. Comparing the accuracy change graphs and loss rate change graphs of GE-DenseNet and DenseNet201, it could be found that the proposed method in this study was effective in training stability, which could accelerate the speed of model convergence and improve the stability of the model, making the network training process more stable. And observing the visualization results of GE-DenseNet and DenseNet201 corresponding feature layers, it could be found that the features were more densely reflected around the pests and diseases after adding the G-ECA Layer structure. From the ablation comparison experiments of the GE-DenseNet model, it could be obtained that the model accuracy increased by 2.27% after the introduction of the Focal Loss function with the G-ECA Layer layer. Comparing the proposed model with the classical NasNet (4@1056), VGG-16 and ResNet50 models, the classification accuracy increased by 6.53%, 4.83% and 3.69%, respectively. Compared with the original DenseNet201, the recognition accuracy of hispa improved 20.32%.

[Conclusions] The experimental results showed that the addition of G-ECA Layer structure enables the model to more accurately capture feature information suitable for rice pest recognition, thus enabling the GE-DenseNet model to achieve more accurate recognition of different rice pest images. This provides reliable technical support for timely pest and disease control, reducing crop yield loss and pesticide use. Future research can lighten the model and reduce its size without significantly reducing the recognition accuracy, so that it can be deployed in UAVs, tractors and various distributed image detection edge devices to facilitate farmers to conduct real-time inspection of farmland and further enhance the intelligence of agricultural production.

Key words: DensetNet201; attention mechanism; pest and disease identification; transfer learning; CNN; Ghost module